

Exploring Preservation Services in a Mediated Peer to Peer Federation



Library of Congress Designing Storage Architectures September 2017

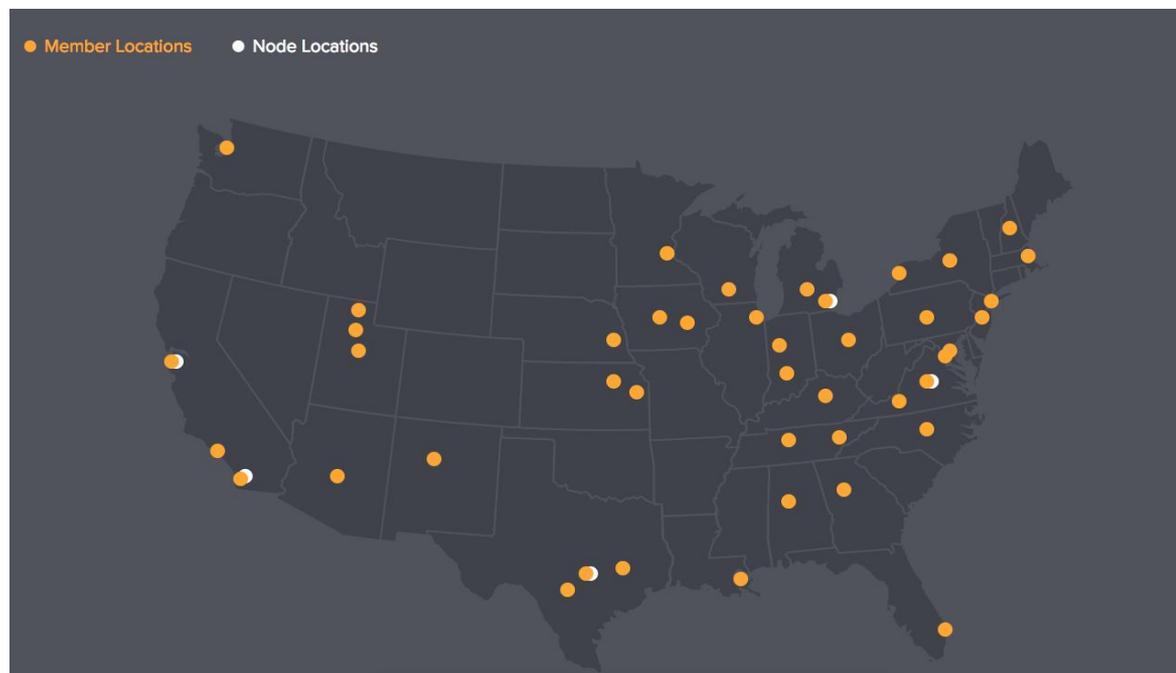
David Pcolar - Technical Officer

dave@dpn.org

<http://www.dpn.org>

What is DPN?

- A membership organization dedicated to the long-term preservation of the scholarly record
- A federation of cooperating repositories (nodes) supporting dark preservation
- A set of services and agreements that, at the time of deposit, commit to ensure that scholarly assets remain available for restoration by the academy



<http://dpn.org/about>

Q: What are the greatest challenges we are facing in implementing preservation solutions?

A: Sustainability and Community Involvement

“The fundamental problem is not storing bits safely for the long term, it is paying to store bits safely for the long term.”

- David Rosenthal

“Traditional business models relied on planning, which in turn assumes either that the future will be like today or that we are good at forecasting. Neither has proved true.”

- Peter Cappelli

Q: What key best practices has your organization discovered or enabled as you implement preservation solutions?

A: Exploring new Technical, Policy, and Economic Models

Community Driven Long Term Preservation

1. Support **new types of services with Peers offering long-term storage**
2. Develop a **pilot project to explore scenarios** for the community.
3. **Explore service options** that support variable numbers of copies.
4. Charge a Working Group to **explore large-scale data movement tools**.
5. Charge a Working Group to **develop financial and governance structures**.

P2P as a Preservation Architecture

“[#DigitalPreservation](#)’s future isn't a central repo, but leveraging distributed assets.”³

DPN’s core Preservation Model continues to be the main service offering.

However...

Members desire to incorporate DPN into a broader institution centered model:

- where DPN’s 3-copy strategy might be overkill
- where institutional requirements drive preservation policy
- where members have in-kind resources that they can commit

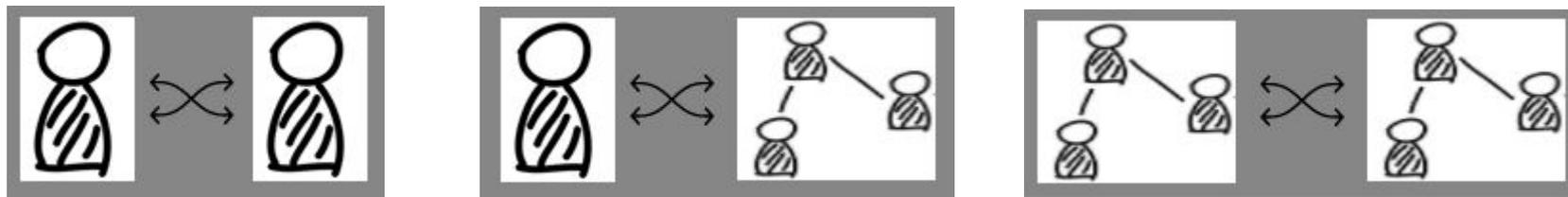
DPN P2P Pilot

Initial Steps

1. **Establish a high level roadmap of the pilot process and goals**
2. **Define distinct areas of exploration**
 - a. Policy
 - b. Membership/Finance
 - c. Technical
3. **Identify and report known constraints**
4. **Pair partners for testing**



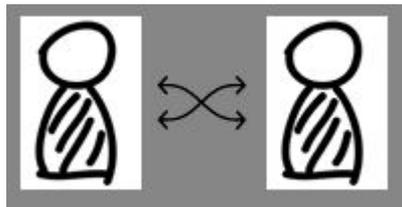
Peer Scenarios



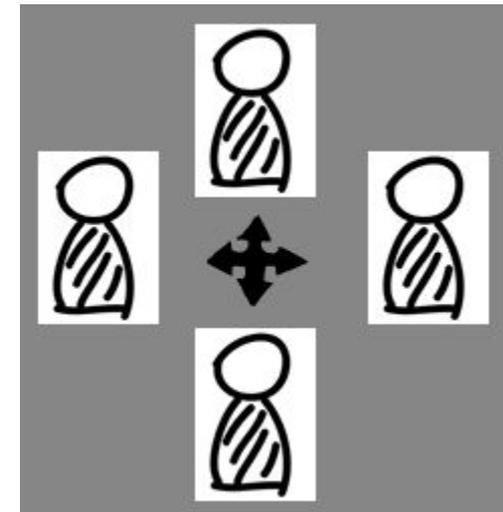
Exploring new transaction models

- Negotiate common transfer and deposit agreements
- Identify DPN's role in the transactions
- Establish workflows for content registration and metadata exchange
- Evaluate sustainability of the models

Preservation Architecture Value Chain



- Emphasizes the importance of collaboration
- Effective partners align and complement skill sets
- Equitable resource commitment and consumption
- Services may not be linear, encouraging many:many relationships
- Trust and commitment are inherent in the relationships



Rethinking Digital Preservation

- Preservation is a game of probabilities
- Selective actions mitigate risk of loss
- Metadata creation, policies, and procedures are foundational elements
- Diverse technical architectures reduce risk

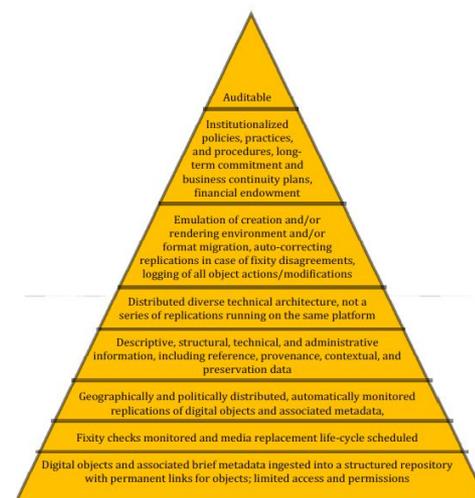


Figure 1: Hierarchy of Digital Preservation Needs [1, 2]

“Sustainable preservation strategies are not built all at once, nor are they static.”

Joint Preservation Code of Ethics

The mission of our shared work is rooted in a set of core values. These core values inform and direct our collective work:

Collaboration - working together and with others to preserve digital assets

Affordability - pricing to sustain and grow services

Availability - access to a variety of services for all, regardless of size or technical ability

Diversity - in our partners, collections, and technical solutions

Stewardship Continuity - by working together to help identify new locations for content that is at risk when one of our organization's stewardship cannot continue.

http://bit.ly/Preservation_codeof_Ethics



Thank You!

dave@dpn.org

P2P References:

- [Long-term Digital Archiving Based on Selection of Repositories Over P2P Networks](#)
- [Peer-to-peer data preservation through storage auctions](#)
- [Network Neutrality \[Paid peering: pricing and adoption incentives\]](#)
- [Internet service provision and content services: paid peering and competition between internet providers](#)
- [KopperCoin – A Distributed File Storage with Financial Incentives](#)
- [P2P commercial digital content exchange](#)
- [Peer to peer data trading to preserve information](#)
- [Creating trading networks of digital archives](#)